PROCEEDINGS OF TDWG

Biodiversity Information Science and Standards 1: e21113
doi: 10.3897/tdwgproceedings.1.21113

OPEN ACCESS

Conference Abstract

Linking Heterogeneous Data in Biodiversity Research

Pamela S Soltis ‡

‡ Florida Museum of Natural History, University of Florida, Gainesville, United States of America

Corresponding author: Pamela S Soltis (psoltis@flmnh.ufl.edu)

Received: 20 Sep 2017 | Published: 20 Sep 2017

Citation: Soltis P (2017) Linking Heterogeneous Data in Biodiversity Research. Proceedings of TDWG 1: e21113. https://doi.org/10.3897/tdwgproceedings.1.21113

Abstract

Emerging cyberinfrastructure and new data sources provide unparalleled opportunities for mobilizing and integrating massive amounts of information from organismal biology, ecology, genomics, climatology, and other disciplines. Key among these data sources is the rapidly growing volume of digitized specimen records from natural history collections. With millions of specimen records currently available online, these data provide excellent information on species distributions, changes in distributions over time, phenology, and a host of traits. Particularly powerful is the integration of phylogenies with specimen data, enabling analyses of phylogenetic diversity in a spatio-temporal context, the evolution of niche space, and more. However, a major impediment is the heterogeneous nature of complex data, and new methods are needed to link these divergent data types. Challenges involve assembly, management, and sharing of data, taxonomic name resolution, the patchy nature of data availability, varying scales of data collection, and data integration.

Through case studies that link and analyze specimen data and related heterogeneous data sources to address a range of evolutionary and ecological problems, we will explore the specific challenges encountered and how these challenges may be overcome. Although many specific hypotheses may be addressed through integrated analyses of linked biodiversity and environmental data, the additional value of such data-enabled science lies in the unanticipated patterns that emerge.

© Soltis P. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords

natural history specimens, heterogeneous data, iDigBio, data integration

Presenting author

Pamela S. Soltis